Release Notes for Platform **LSF**®
Version 5.0

October 3 2002
Platform Computing
Comments to: doc@platform.com

Contents

Platform™

What's New in the Platform LSF Version 5.0

Platform **LSF**® Version 5.0 ("LSF") introduces the following new features:

Robustness, performance and scalability

Performance and scalability are key requirements for Enterprise deployments. New Platform LSF Version 5.0 architecture enhances performance and scalability:

- Manage more than 500,000 active jobs in the system in clusters of more than 1500 hosts
- For clusters with multi-CPU hosts, you can build clusters of more than 10,000 CPUs
- Fairshare scheduling achieves nearly 100% resource utilization; job throughput improved by 30%
- Response time for job submission, job control, and information query commands (e.g., `bsub`, `bjobs`, `bstop`) improved by 20%
- Manage more than 100 clusters with Platform MultiCluster

Scheduler plugin SDK

Write customized scheduler plugin modules to give you more flexibility and control over job scheduling. Enable your own custom scheduling policies by configuring your modules in `lsb.modules`.

The directory

`LSF_TOP/5.0/misc/examples/external_plugin/`

contains sample plugin code. See *Using the Platform LSF SDK* for more detailed information about writing, building, and configuring your own custom scheduler plugin.

Platform Workload Management Web GUI

The Platform Workload Management Web GUI allows users to submit, control, and monitor their work over the Internet. Administrators can use the Web GUI to manage the cluster, including queues and hosts. The benefit is ease of use, and convenient access to distributed computing from anywhere and anytime.

LSF Web Service Broker provides an open, SOAP-based XML interface that allows users to write applications to access the LSF capabilities from anywhere any time.

The UI is customizable to meet any site-specific requirements, and the XML interface allows easy application integrations as it is standards-based.

Policy management

Advance reservation

With advance reservation you can now ensure that a set of hosts is exclusively dedicated to specific jobs, without having to worry about other jobs using those hosts. The result is higher throughput and reliability for critical jobs.

Advance reservations guarantee access to specific hosts during specified times. An advance reservation is essentially a lock on a number of processors. The reservation is active only within the time frame specified, and any given host may have several reservations in place, some of which may be active at the same time.

During the time the reservation is active, only users or groups associated with the reservation have access to run jobs on the reserved hosts. Reservations can also be created for system maintenance. If a system reservation is active, no other jobs can use the reserved hosts.

Resource allocation limits

Manage complex resource allocation policies with enhanced resource allocation limits. You now have complete flexibility in defining how resources are used in your cluster:

- All limits are configured in a single file, `lsb.resources`
- You can set and combine resource limits based on users, projects, queues, and hosts for:
  - Job slots per host
  - Licenses
  - Memory (in MB or as a percentage per host)
  - Swap space (in MB or as a percentage per host)
  - Tmp space (in MB or as a percentage per host)
  - Any shared resource that can be reserved
- You can set cluster-wide limits for all limits or combine resource limits and consumers
- You can set limits for all consumers of a specified resource or on a per-consumer basis
- You can set limits for slots and any resource that can be reserved for:
  - Multiple queues
  - Users and user groups
  - Hosts and host groups
  - Projects
  - Jobs running on different hosts from the same queue
  - Jobs from different queues running on the same host

Fairshare across queues

Fairshare scheduling is enhanced to give a site more control managing the processing power of the LSF cluster among users and groups. Cross-queue fairshare provides fair access to resources for all jobs across multiple queues, to ensure that the right amount of resources are allocated to the right users. You can now manage multiple fairshare policies that map to your business needs while maximizing resource usage.

You can define a fairshare policy that applies to several queues at the same time with the parameter FAIRSHARE in `lsb.queues`. You define the fairshare policy in a *master queue* and list *slave queues* to which the same fairshare policy applies; slave queues inherit the same fairshare policy as your master queue. A user has the same priority across the master and slave queues.

In this way, if a user submits jobs to different queues, user priority is calculated by taking into account all the jobs the user has submitted across the defined queues.

Chunk job enhancement

When CHUNK_JOB_DURATION is set in `lsb.params`, the CPU limit or run limit set in the queue (CPULIMIT or RUNLMIT) or specified at job submission (`-c` or `-W bsub` options) must be less than or equal to CHUNK_JOB_DURATION for jobs to be considered for job chunking.

Memory reservation for pending jobs

This enhancement prevents starvation of jobs requiring a large amount of memory, ensuring that memory-intensive jobs are given their share of resources.

By default, the `rusage` string reserves resources for running jobs. Because resources are not reserved for pending jobs, some memory-intensive jobs could be pending indefinitely because smaller jobs take the resources immediately before the larger jobs can start running. The more memory a job requires, the worse the problem is.

Memory reservation for pending jobs reserves memory as it becomes available, until the total required memory specified on the `rusage` string is accumulated and the job can start. Use memory reservation for pending jobs if memory-intensive jobs often compete for memory with smaller jobs in your cluster.

Parallel processes on homogeneous hosts

This enhancement describes how to use the `same` section in the resource requirement string to run parallel processes on only one type or model of host. You can also use a custom resource to define the criteria for homogeneous hosts.

For example, you can specify to run parallel processes only on the same host type and model, or only on hosts that belong to the same high-speed connection group.

Ease of administration

Suspending and resuming jobs

This enhancement gives administrators more control over all jobs in the system to allow them to properly manage their cluster without interference by user actions. Administrators can suspend jobs to do system troubleshooting or maintenance, or stop runaway jobs or jobs with errors, without users overriding these actions by resuming the suspended jobs.

Jobs that are suspended by the administrator can only be resumed by the administrator or root; users do not have permission to resume a job suspended by another user or the administrator. Administrators can resume jobs suspended by users or administrators.

To enable users to resume their own jobs that have been suspended by administrators, set ENABLE_USER_RESUME=Y in `lsb.params`.

Platform MultiCluster

Platform MultiCluster is Platform's Enterprise Grid solution, offering improved scalability, transparency, security, and ease of administration. LSF cluster administrators have full control of optimal resource sharing policies across remote sites.

For additional details, see the *Platform MultiCluster Guide*.

Platform MultiCluster provides the following enhancements:

- Enhanced single-system image for better cluster management
- Improved transparency across clusters
- MultiCluster now supports the following features
  - FCFS scheduling order preservation
  - Fairshare scheduling
  - Preemption scheduling
  - CPU reservation and backfill
  - Memory reservation and backfill
  - Advance reservation
  - Queue switch support
  - Job pending and suspending reasons
  - Limits on any resource that can be reserved based on users, projects, queues, and hosts
  - Chunk jobs
  - Exclusive jobs
  - Full job control
- Different authentication methods can be used by different clusters
- Daemon communication over non-privileged ports
- Non-uniform name spaces and file transfer

Windows enhancements

Platform LSF on Windows offers easier installation, configuration, deployment, and administration.

Mixed UNIX and Windows clusters no longer need a shared file system

It is now possible to install a mixed cluster without having a shared file system across UNIX and Windows hosts.

In this type of configuration:

- An `lsf.conf` file is created in the Windows shared directory.
- LSF_MASTER_LIST is defined in the UNIX `lsf.conf`.
- An encrypted password file `passwd.lsfuser` is created in the Windows shared directory.

To install a mixed cluster without a shared file system between UNIX and Windows hosts, install a UNIX cluster, then add Windows hosts to the cluster. What you enter at installation prompts determines whether you have a shared file system or not.

External checkpoint support for Windows 2000 and Windows XP (echkpt and erestart)

LSF Version 5.0 supports application-level checkpointing on Windows 2000 and Windows XP hosts.

Checkpointing is not supported on Windows NT hosts. User-level and kernel-level checkpointing are not supported on Windows.

Windows XP support

LSF is now supported on Windows XP.

Upgrade and Compatibility Notes

UPGRADE document

To upgrade from LSF Version 4.*x* to LSF Version 5.0, follow the steps in `upgrade.html`.

Compatibility between LSF 4.x and Version 5.0

Platform LSF

- You must upgrade the LSF master hosts in your cluster to Version 5.0.

  LSF 4.*x* servers are compatible with Version 5.0 master hosts. All LSF 4.x features are supported by 5.0 master hosts except:

  - 4.*x* `bhist` cannot work in an LSF Version 5.0 cluster.

    Only the Version 5.0 `bhist` command can access the encrypted `lsb.events`. See `upgrade.html` for details.

  - Jobs that are suspended by the administrator can only be resumed by the administrator or root.

    To enable users to resume their own jobs that have been suspended by administrators, set ENABLE_USER_RESUME=Y in lsb.params.

  If you need 5.0 features, upgrade all hosts in your cluster into 5.0.

- The following configuration parameters are no longer needed as a result of improved fairshare scheduling performance:
  - ENABLE_AUTOADJUST_AT_PERCENT
  - AUTOADJUST_AT_PERCENT
  - AUTOADJUST_AT_NUM_PEND

Platform MultiCluster

- You must upgrade the LSF master hosts in all clusters to Version 5.0.
- The MultiCluster job forwarding model in Version 5.0 is based on MultiCluster version 4.*x*, with many enhancements. MultiCluster queues configured for version 4.*x* can also work with Version 5.0. A general rule for the MultiCluster job forwarding model is that a MultiCluster job is affected by the policies of the execution queue, not the submission queue.

Platform LSF Version 5.0 API

The Platform LSF Version 5.0 API is fully compatible with the LSF Version 4.*x* API. An application linked with the LSF Version 4.*x* library will run under LSF Version 5.0 without relinking.

An existing LSF application linked with LSF Version 4.*x* can be relinked or recompiled with LSF Version 5.0, and be able to run under LSF Version 5.0 without changing any code.

`lsb.events` is encrypted in LSF Version 5.0. If your application uses the `lsb_geteventrec()` and `lsb_puteventrec()`APIs, see `upgrade.html` for details.

New files added to installation

The following new files have been added to the Platform LSF installation:

- `LSF_BINDIR/brsvadd`
- `LSF_BINDIR/brsvdel`
- `LSF_BINDIR/brsvs`
- `LSF_SERVERDIR/mbschd`
- `LSF_LIBDIR/schmod_advrsv.so`
- `LSF_LIBDIR/schmod_default.so`
- `LSF_LIBDIR/schmod_fairshare.so`
- `LSF_LIBDIR/schmod_limit.so`
- `LSF_LIBDIR/schmod_mc.so`
- `LSF_LIBDIR/schmod_parallel.so`
- `LSF_LIBDIR/schmod_preemption.so`
- `LSF_LIBDIR/schmod_topology.so` (for SGI systems running IRIX 6.5.8 and up)
- `LSB_CONFDIR/`*`cluster_name`*`/configdir/lsb.resources`
- `LSB_CONFDIR/`*`cluster_name`*`/configdir/lsb.modules`

If your installation uses symbolic links to other files in these directories, you must manually create links to these new files.

Learning About Platform LSF Version 5.0

Finding Platform LSF Version 5.0 information

Information about Platform LSF Version 5.0 is available online from the following sources:

- "World Wide Web and FTP"
- "README, release notes, and UPGRADE"
- "Platform LSF documentation"

World Wide Web and FTP

The latest information about all other supported releases of Platform LSF is available on the Platform Web site at `www.platform.com`.

If you have problems accessing the Platform Web site or the Platform FTP site, send email to `support@platform.com`

README, release notes, and UPGRADE

Before downloading and installing LSF, be sure to read the files named `readme.html` and `release_notes.html`.

To upgrade from LSF Version 4.*x* to LSF Version 5.0, follow the steps in `upgrade.html`.

Platform LSF documentation

The LSF documentation directory `/distrib/5.0/docs/` contains the LSF documentation set in HTML and PDF format.

Download or view LSF documentation online

View or download the LSF documentation in HTML or PDF format:

- LSF Documentation CD (Disc 3) included in this package
  The file `readme.txt` describes the contents of the LSF Documentation CD
- Platform Computing Web site:
  `www.platform.com/lsf_docs`
- Platform Computing FTP site (`ftp.platform.com`):
  `/distrib/5.0/docs/`

Online manuals

Title	PDF	HTML	
		.zip	.tar.Z
Installing Platform LSF on UNIX (`lsfinstall`)	70 KB	240 KB	290 KB
Running Jobs with Platform LSF	830 KB	1,300 KB	1,300 KB
Platform LSF Administrator's Primer	575 KB	1,300 KB	1,300 KB
Administering Platform LSF	4,410 KB	1,300 KB	1,300 KB
Platform LSF Reference	3,770 KB	360 KB	420 KB

Title	PDF	HTML	
		.zip	.tar.Z
Platform LSF Quick Reference	90 KB	—	—
Platform LSF on Windows	1,150 KB	240 KB	290 KB
Licensing Platform LSF	460 KB	—	—
Platform MultiCluster Guide	780 KB	240 KB	290 KB
Using the Platform LSF SDK	1,060 KB	240 KB	290 KB
Using Platform LSF Parallel	470 KB	110 KB	120 KB
Complete Platform LSF Version 5.0 HTML Doc Set		4,000 KB	4,650 KB

LSF integrations

Title	PDF
Using LSF with Maya	80 KB
Using LSF with Unicenter TNG	110 KB
Using LSF with ClearCase	80 KB
Using LSF with Condor Checkpointing	80 KB
Using LSF with IRIX cpusets	160 KB
Frame Arrays	105 KB
Using LSF with FLUENT	85 KB
Using LSF with IBM-SP2	65 KB
Using LSF with Kerberos V	100 KB
Using LSF with LAM/MPI	100 KB
Using LSF with MPICH-GM	110 KB
Using LSF with SNMP	130 KB

Documentation for these and other LSF integrations is also provided with the integration distribution packages.

Other LSF environments

Title	PDF
Installing LSF on AFS	135 KB
Installing LSF on DCE/DFS	80 KB
Using LSF with Solaris Processor Sets	75 KB
Using LSF with IRIX 6 Processor Sets	75 KB

Companion products

Title	PDF
LSF Make	75 KB

Getting Technical Support

Contacting Platform

Contact Platform Computing or your LSF vendor for technical support. Use one of the following to contact Platform technical support:

Email support@platform.com

World Wide Web www.platform.com

Phone
- North America: +1 905 948 4297
- Europe: +44 1256 370 530
- Asia: +86 10 6238 1125

Toll-free Phone 1-877-444-4LSF (+1 877 444 4573)

Mail
Platform Support
Platform Computing Corporation
3760 14th Avenue
Markham, Ontario
Canada L3R 3T7

When contacting Platform, please include the full name of your company.

We'd like to hear from you

If you find an error in any Platform documentation, or you have a suggestion for improving it, please let us know:

Email doc@platform.com

Mail
Information Development
Platform Computing Corporation
3760 14th Avenue
Markham, Ontario
Canada L3R 3T7

Be sure to tell us:

- The title of the manual you are commenting on
- The version of the product you are using
- The format of the manual (HTML or PDF)

Copyright